DOCUMENT RESUME

ED 476 921 TM 034 972

AUTHOR	Kobrin, Jennifer L.; Milewski, Glenn B.; Everson, Howard; Zhou, Ying
TITLE	An Investigation of School-Level Factors for Students with Discrepant High School GPA and SAT Scores.
PUB DATE	2003-04-00
NOTE	29p.; Paper presented at the Annual Meeting of the National Council on Measurement in Education (Chicago, IL, April 22-24, 2003).
PUB TYPE	Reports - Research (143) -- Speeches/Meeting Papers (150)
EDRS PRICE	EDRS Price MF01/PC02 Plus Postage.
DESCRIPTORS	*College Entrance Examinations; Curriculum; *Grade Point Average; *High School Students; High Schools; Mathematics Tests; Prediction; Scores; Sex Differences; Socioeconomic Status; Student Characteristics; Verbal Tests
IDENTIFIERS	Discrepancy Analysis; *Scholastic Assessment Tests

ABSTRACT

This study used hierarchical linear modeling (HLM) to examine student- and school-level predictors of the discrepancy between students' standardized high school grade point average (HSGPA) and standardized total Scholastic Assessment Test (SAT) scores. At the student level, academic curriculum intensity, socioeconomic status (SES), the difference between a student's SAT mathematics and verbal scores (SATM-V), and gender were used to predict the HSGPA-SAT discrepancy within each school. Four factor scores (economic advantage, school size, computer technology, and school resources) based on a principal components analysis of 13 school-level variables were used to predict variation in the intercepts and slopes across schools. Data from the College Board for 18,674 students were used. All of the student-level variables except for curriculum intensity were significant predictors of discrepancy scores. Level-one intercepts as well as slopes for gender varied significantly across schools; the slopes for the other student-level variables did not. The school-level factor scores for economic advantage and school size significantly predicted a school's average discrepancy score (or level-one intercept), and the economic advantage factor also predicted a school's slope for gender. While several of the student-level variables were significant predictors of discrepancy scores, a substantial amount of the variance remained unexplained. This suggests that other variables not examined in this study are important predictors of the discrepancy between high school grades and SAT scores. (Contains 8 tables, 1 figure, and 23 references.) (Author/SLD)

Reproductions supplied by EDRS are the best that can be made
from the original document.

ED 476 921

Running head: SCHOOL-LEVEL FACTORS FOR DISCREPANT SCORES

An Investigation of School-Level Factors for Students with Discrepant High School GPA and SAT® Scores

Jennifer L. Kobrin, Glenn B. Milewski, and Howard Everson

The College Board

Ying Zhou

Fordham University

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☑ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY
J. L. Kobrin
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
1

TM034972

Paper presented at the annual meeting of the National Council on Measurement in Education

Chicago, Illinois, April 22, 2003

BEST COPY AVAILABLE

Abstract

This study used hierarchical linear modeling (HLM) to examine student- and school-level predictors of the discrepancy between students' standardized high school grade-point average (HSGPA) and standardized total SAT scores. At the student level, academic curriculum intensity, socio-economic status (SES), the difference between a student's SAT Math and Verbal score (SATM-V), and gender were used to predict the HSGPA-SAT discrepancy within each school. Four factor scores (economic advantage, school size, computer technology, and school resources) based on a principal components analysis of 13 school-level variables were used to predict variation in the intercepts and slopes across schools. All of the student-level variables except for curriculum intensity were significant predictors of discrepancy scores. Level-one intercepts as well as slopes for gender varied significantly across schools; the slopes for the other student-level variables did not. The school-level factor scores for economic advantage and school size significantly predicted a school's average discrepancy score (or level-one intercept), and the economic advantage factor also predicted a school's slope for gender. While several of the student-level variables were significant predictors of discrepancy scores, a substantial amount of the variance remained unexplained. This suggests that other variables not examined in this study are important predictors of the discrepancy between high school grades and SAT scores.

An Investigation of School-Level Factors for Students with Discrepant High School GPA and SAT® Scores

Most colleges and universities in the United States use both high school grades and SAT® I: Reasoning Test scores (hereafter referred to as the SAT) when making admissions decisions (Breland, Maxey, Gernand, Cumming, Trapani, 2002). These institutions rely heavily on SAT scores and grades because they provide non-redundant information about students' likelihood of academic success in college (Koretz & Berends, 2001). The correlation between these two measures is about .47, suggesting that high school grades and the SAT measure related constructs. Yet, sometimes a student's high school grade point average (HSGPA) and SAT score are inconsistent and contradictory. That is, they present a high HSGPA and low SAT score, or vice versa. These applicants often present a challenge to the admission staff who decide whether or not to admit them to their college or university. To help better understand these applicants, we have been conducting research into who they are and how they do in college.

Recently, for example, Kobrin and Milewski (2002) examined predictions of first-year college performance for students whose SAT score and high school grades were discrepant and compared them with students whose SAT score and high school grades were more consonant. Their data were from 48,410 students who entered as college freshmen in 1994 or 1995 at 23 different institutions (see Bridgeman, McCamley-Jenkins, & Ervin, 2000 for details on the sample). Kobrin and Milewski replicated an earlier study by Baydar (1990) that was based on three colleges.

Kobrin and Milewski (2002) computed standardized scores[1] for the combined SAT (Verbal and Math) and self-reported, cumulative high school grades. They created three groups

[1] Standardized scores are expressed in standard deviation units and provide a measure of an individuals' relative standing in a group (Vogt, 1999).

of students based on a comparison of standardized HSGPA and SAT scores; these groups were labeled (a) non-discrepant scores (NDS), (b) HSGPA discrepant scores (HSD) and (c) SAT discrepant scores (SATD). Students in the NDS group had a standardized SAT score that was within one standard deviation of their HSGPA score. Students in the HSD group had a standardized SAT score that was more than one standard deviation below, or about 34% lower than their HSGPA. Students in the SATD group had a standardized SAT score that was more than one standard deviation above, or about 34% percent higher than their HSGPA.

Kobrin and Milewski (2002) found that 68% of students were categorized as NDS, 16.2% (N = 7,837) as HSD, and 15.8% (N = 7,653) as SATD. Female and ethnic minority students (of Asian, Black or African American, or Hispanic descent) were much more heavily represented in the HSD group than in either the NDS or SATD groups. In addition, a higher percentage of students in the HSD group spoke languages other than English, were not U.S. citizens or nationals, and had relatively lower family income. Because these students' SAT scores are not consistent with their otherwise high performance in high school, some infer that the SAT may not be measuring well the reasoning skills of these students, and may even disadvantage them in the college admission process.

While it is the case that students in the HSD group had higher high school grades than students in the NDS or SATD groups, Kobrin and Milewski (2002) reported that these students did not have higher first-year college grades (FGPA) than the other students. As shown in Table 1, both SAT and HSGPA accounted for a smaller amount of the variance of first-year college GPA for the HSD group than for the other two groups, as measured by R-square. In addition, the SAT had greater incremental validity over HSGPA for the HSD group than for the other two groups. Kobrin and Milewski also found that FGPA was overpredicted to a greater extent in the HSD group than in the other two groups.

Table 1
Regression of SAT and HSGPA on FGPA

Predictor	Statistic	NDS (N=32,916)	HSD (N=7,837)	SATD (N=7,653)
SAT Verbal	R-square	.167	.093	.182
	MSE	(.463)	(.475)	(.563)
SAT Math	R-square	.166	.106	.152
	MSE	(.464)	(.468)	(.584)
SAT Total	R-square	.209	.144	.205
	MSE	(.446)	(.449)	(.547)
HSGPA	R-square	.213	.127	.215
	MSE	(.438)	(.457)	(.540)
HSGPA + SAT (V&M)	R-square	.232	.150	.225
	MSE	(.428)	(.445)	(.533)
SAT Incremental Validity	R-square change	.019	.023	.010
	MSE change	(.010)	(.012)	(.007)

Note. From Kobrin & Milewski (2002). All R-square values significant at the .01 level.

These findings are neither new nor surprising. In a study of the trends in predictive validity of the SAT and high school rank from the mid-1970's to the mid-1980's, Willingham, Lewis, Morgan, and Ramist (1990) found that the SAT provided a slightly better prediction of FGPA for students in the lower third of the class than did high school rank, and noted that this is contrary to the usual assumption that performance in high school gives a much better indication of academic promise than does the SAT for students who are in the gray area of an institution's applicant selection range.

The SATD group, in contrast, may comprise students who are academically able but not motivated to put forth the effort to achieve good grades in high school, and/or students from poorer schools who do not have access to rigorous academic courses and thus do not perform well on tests of developed ability. These students are sometimes called "diamonds in the rough," i.e., students with talent and promise who do not have exceptional high school grades. These

students are an interesting group who are not well understood. As shown in Table 1, the predictive validity of the SAT and HSGPA is similar for students in the SATD group and students with non-discrepant scores. However the mean square error[2] for the SATD group was higher than either of the other two groups, indicating that regression-based models may overestimate first-year college GPAs for this group.

The reasons for discrepancies in SAT scores and HSGPA are multifaceted and complex. Factors at the school level, including the phenomenon of grade inflation and differences in high school teachers' grading practices probably contribute to this discrepancy. Individual-level factors, such as family background and motivation also are likely to play a role in producing these discrepancies. Oakes (1990) noted three domains of influence on students' academic achievement: cognitive abilities and attitudes of individual students, schooling factors and opportunities, and societal factors. In the current study, multi-level modeling techniques were employed to examine the role of individual and school-level factors in the discrepancy between HSGPA and SAT scores. Before describing the results of our study, a brief review of the literature related to these factors is presented below.

---

[2] From Kobrin and Milewski (2002): Since the R-square is a function of the variance in the dependent variable (R-Square = 1 – Mean square error (*MSE*)/ σ_y^2), sampling methods that produce a *real* decrease in the variance of the dependent variable result in a restriction in range and consequently, an underestimation of R-square (Nunnaly, 1978). Two of the sampling methods applied in the current study resulted in a restriction in range. First, since not all of the students that submitted HSGPA and SAT scores to the 23 schools considered in the analysis enrolled and earned an FGPA score, the variance on FGPA is smaller than it would have been had all the applicants completed a year of study. Second, the selection criteria for the three discrepant score groups resulted in a decrease in the variance of FGPA across groups. Normally, underestimation of R-square due to restriction in range would be corrected for using the standard Pearson-Lawley procedure (Gulliksen, 1950). However, since this procedure is based on several important assumptions that were not met in the current study (e. g. homogeneity of regression), it was not appropriate to perform this correction. In an effort to present a more accurate representation of the relationship between SAT, HSGPA and FGPA in each group, the Mean square error (*MSE*) was reported. The *MSE* is "a measure of the degree of variability of the points around a regression line" (Vogt, 1993) and can be used as an indicator of the strength of the relationship between predictors and criterion. The smaller the *MSE*, the stronger the relationship between predictor(s) and criterion.

BEST COPY AVAILABLE

Literature Review

Grade inflation refers to a situation in which increasing grades are not commensurately reflected in increased academic achievement (Zirkel, 1999, cited in Mulvenon & Thorn, 2002). Grade inflation is usually measured by comparing the increase in students' grades over time with the increase in other measures of ability (e.g., SAT or achievement test scores). Evidence of grade inflation has been documented in the literature (Bejar & Blew, 1981; Cizek, 2000; Ziomek & Svec, 1995), although the reasons for the presence of grade inflation are difficult to pinpoint (Cizek, 2000).

To further complicate matters, there is evidence that grade inflation has been more apparent with some groups than with others. Koretz and Berends (2001) examined ten years of high school grade data (1982 to 1992) from the High School and Beyond (HSB) and National Educational Longitudinal Study (NELS) surveys. They found a general increase in the mean high school GPA from 2.56 to 2.63 over the ten years. The percentage of students receiving a GPA of 3.0 and higher increased from 42 to 46.2, and the percentage with GPA of 3.3 or higher increased from 27.7 to 30.8. When examining the changes in GPAs by ethnic group, only Hispanic students showed a substantially larger increase in mean GPA compared to the other ethnic groups. More substantial differences appeared across income categories. The mean increase in the highest income category was nearly three times as large as the overall mean increase. The change in GPA also varied substantially depending on school location. The mean GPA of students in rural and suburban schools increased by only .04, while students in urban schools increased by .22.

Similarly, the U.S. Department of Education's Office of Educational Research and Improvement (1994) found that eighth grade "A" students in high poverty schools (those where more than 75 percent of students receive free or reduced price lunch) received lower scores on

the NELS:88 tests, on average, than their counterparts in the more affluent schools. Students in high poverty schools who received mostly A's in English got about the same reading score as did the "C" and "D" students in the most affluent schools. In math, the "A" students in the high poverty schools most closely resembled the "D" students in the most affluent schools. Clearly, this research suggests that high school grades are not equivalent across schools, which may explain why SAT and other test scores are sometimes discrepant from high school grades.

The non-standard grading practice of teachers across schools is another likely cause of discrepant SAT scores and HSGPA. In evaluating students, teachers consider a number of attributes in deciding a student's grades, such as achievement, motivation/effort, and ability (Pilcher, 1994). Measurement specialists recommend that only achievement be considered when assigning a grade, so that it is a unidimensional measure and instruments or tests can be developed to measure it objectively. However, teachers have been found not to follow these recommendations, whether or not they have been trained in classroom assessment procedures (Stiggins, Frisbie, & Griswold, 1989).

Many teachers assign grades based on the extent to which content was mastered for a student with a certain ability level. Therefore, it is possible for students to achieve more than a grade indicates, and it is also possible for students to receive a grade that is higher than their achievement level (test scores). In Pilcher's (1994) case study, three students who were perceived to have high ability in their classroom but did not consistently put forth effort were penalized by their teachers' grading practices. The other three students who were perceived to have lower ability in their classroom and completed their assignments were awarded when they applied effort.

The research suggests that students with a high HSGPA in the presence of low SAT scores do not do any better in college than students with lower HSGPA scores but higher SAT

scores. Therefore, the SAT may be a more accurate predictor than HSGPA for the students in the HSD group. This finding is problematic for college admissions staff who rely on these two pieces of information to help make admissions decisions, and often assign the same weight to HSGPA and SAT for all applicants, or give more weight to HSGPA. Sometimes colleges collect information about applicants' high schools and courses taken and use this information to help interpret grades provided on a transcript. For example, the GPAs of students who have taken more difficult courses are sometimes given more weight than those of students taking less difficult courses. Yet, often this process is not informed by any concrete understanding of what types of students from what types of schools are more likely to present discrepant information. A more thorough understanding of the student and school-level factors associated with students who have discrepant scores would enable college admission officers to make better decisions regarding admission.

As noted earlier, the purpose of this study is to examine both the student and school factors that predict discrepancies between high school grades and SAT scores, in order to gain a better understanding of this phenomenon. Previous research by Everson and Millsap (1999) that employed multi-level models found that school-level variables, i.e., the percentages of free-lunch eligible students and the proportion of non-white students in a school predicted students' SAT verbal and math scores. The current study will attempt to explore whether these and other student and school characteristics predict discrepancies between high school grades and SAT scores.

Method

Data Sources

Student-level data were obtained from a database compiled by The College Board and Educational Testing Service and includes demographic information (gender, ethnicity, first language and best language, parents' education), highest combined (verbal + math) SAT score, and self-reported high school grade point average for 18,674 students from 949 high schools in the United States. These students completed their first year of college during the 1995-1996 academic year. Only high schools with at least ten students were included in the analysis[3], and the average school sample size was 20 students.

High School-level data were obtained from the Quality Education Database (QED) published in 2000, a comprehensive national database rich with information on school demographics, resources and staffing, faculty, and both academic and nonacademic programs and services. The QED was merged with the student-level data through unique high school numbers (i.e., Attending Institution [AI] Codes) assigned by the College Board. The QED data were only available starting in 2000, and student-level data were only available for 1995; therefore before matching the data, it was necessary to ensure that school-level data did not change substantially over time. Otherwise, it would be inappropriate to pursue examining relationships between these two levels of data. In order to do this, we consulted the National Center for Education Statistics (NCES) website, which is another source of high school-level data, albeit not as comprehensive as the QED database. A few variables measured by NCES were also measured by QED, specifically, the percentage of students eligible for free or reduced lunch, and the percentage of minority (non-White) students. To compare these variables across

the two databases, 100 schools were randomly selected from the QED database, and these schools were matched with the 1995 NCES data by using a combination of school number within state and school name, with a 97% success rate. Of these 97 cases, 40 had valid values for the percent free lunch, and 42 had valid values for the percent of minority students. The mean differences between the 1995 NCES and 2000 QED data were very small, .02 for the percent of minority students and .03 for the percent of students eligible for free lunch. These results suggest that it is indeed appropriate to use the 2000 school-level data with 1995 student-level data.

The demographic and academic characteristics of the sample used in this study and those of the population of college-bound seniors in 1995 are shown in Table 2. It is shown that this sample over-represents Asian and White students and under-represents African American and Hispanic students. Furthermore, the mean SAT scores of the sample are much higher than the college-bound population in 1995. The differences between the sample and population should be taken into account when generalizing the findings of this study.

Procedures

A Hierarchical Linear Model (HLM) was used to model SAT and HSGPA discrepancy as a function of student characteristics in each school, and to explain variation among schools in the intercept and slope of the regression equation (if any) in terms of school characteristics. The structure of the data is hierarchical because students are nested within different high schools. The advantage of considering the hierarchical structure of these data is that student behavior is often affected by school characteristics. The computer software HLM 5 (Raudenbush, Bryk, Cheong, & Congdon, 2000) was used to analyze the data.

---

[3] Only the schools with at least ten students were included because schools with "small sample size can easily become an outlier or leverage point because of the instability associated with the limited amount of information

Table 2
Characteristics of the Sample and Population
of 1995 College Bound Seniors

	Sample	1995 CB Seniors
Gender		
Females	52.1%	53.6%
Males	47.9%	46.4%
Race/Ethnicity		
American Indian	0.6%	0.8%
African American	4.0%	9.1%
Asian	11.1%	7.1%
Hispanic	3.9%	7.0%
White	71.0%	59.1%
Other	2.0%	2.2%
First Language		
English only	85.6%	71.1%
English and another	4.8%	7.7%
Another language	1.4%	7.1%
Mean SAT Total	1158	910
Mean SAT Verbal	572	428
Mean SAT Math	584	482
Mean HSGPA	3.5	N/A
Mean FGPA	2.8	N/A

Note. Total N for the sample = 18,674; for the population = 1,140,129

Student-Level Variables. At the first level, academic curriculum intensity (INTENSE), socio-economic status (SES), the discrepancy between a student's SAT Math and Verbal score (SATM-V), and gender were used to predict discrepancy scores within each school. The dependent variable at this level of analysis was the discrepancy between a student's combined SAT score (Verbal + Math) and their self-reported high school grade point average (HSGPA). Discrepancy scores were calculated by subtracting each student's standardized HSGPA from their standardized SAT score (i.e., SAT – HSGPA) resulting in a score in standard deviation units. A score of zero indicates that there is no discrepancy between the student's SAT score and HSGPA, a positive score shows that the student's SAT is higher than their HSGPA, and a negative score shows that a student's HSGPA is higher than their SAT. A score of –2, for example, indicates that a student's HSGPA score is two standard deviations above their SAT

associated with that unit" (Bryk & Raudenbush, 1992, p. 92).

score. The discrepancy scores ranged from –4.32 to 5.02 with a mean of zero and a standard deviation of one.

A variable of academic curriculum intensity modeled after Adelman (1999) was created to represent the rigor of students' high school curriculum. In his study of what contributes most to long-term bachelor's degree completion, Adelman found that the most important variable was Academic Resources – a composite measure of the academic content and performance of the student from secondary school into higher education (composite of high school curriculum, test scores, and class rank). This measure is dominated by the intensity and quality of secondary school curriculum. Adelman's academic intensity variable was constructed to include Carnegie units in six academic areas (English, math, laboratory science and total science, history, social studies, and foreign languages), and also accounted for highest math studied, remedial work in English and math, and advanced placement. Adelman found that the impact of high school curriculum on bachelor's degree completion was far more pronounced for African-American and Latino students than for students in other racial/ethnic groups.

The academic curriculum intensity variable used in the current study (INTENSE) was constructed based on students' responses to the Student Descriptive Questionnaire (SDQ), which was completed when they registered to take the SAT. The SDQ includes questions on the number and types of courses taken in high school, including Advanced Placement (AP) and honors courses. A student earned one point for each of the following courses that he/she reported taking during high school: English, geometry, algebra, trigonometry, precalculus, calculus, computer math/ computer science, other math, geology, other science, social science, and foreign language. A student earned two points for taking biology, chemistry, and physics. Points were assigned regardless of whether the course was regular, honors, or AP. An additional point was given to students who had taken one AP course and two points were given to students

taking more than one AP course; students not taking any AP courses were not given any additional points. Finally, students whose highest math course taken was trigonometry, precalculus, or calculus earned one extra point, students who took algebra as their highest math course did not earn any points, and students who did not reach algebra had one point deducted from their score. The INTENSE variable ranged from -1 to 21 with a mean of 13.35 and a standard deviation of 4.44.

A variable labeled socio-economic status (SES) was created for each student, and was based on a principal components analysis of students' self-reported family income, mother's education, and father's education. The principal components analysis revealed a one-factor solution with loadings of .74, .85, and .81 for the three variables, respectively. Factor scores representing SES were created using the Bartlett method[4]. These scores ranged from –2.9 to 1.7 with a mean of zero and a standard deviation of one.

A variable labeled SATM-V was created to represent the difference between a student's SAT Math and SAT Verbal score (SATM – SATV). Because the SAT-HSGPA discrepancy scores were based on total SAT scores, and these discrepancy scores may vary as a function of whether the student scored higher in verbal or math, this variable was created to account for this information. A positive score indicates that the student scored higher on SAT-Math than SAT-Verbal, and a negative score indicated that the student scored higher on SAT-Verbal than SAT-Math. The SATM-V scores ranged from –300 to 440 with a mean of 8.2 and a standard deviation of 79.5.

School-Level Variables. To reduce the number of independent variables at the school level, a principal components analysis was performed on 13 school-level variables from the QED database. Table 3 provides a description of the school-level variables that were analyzed.

Principal components analysis was selected as the extraction method because it analyzes all of the variance in the observed variables. The Varimax procedure was used to rotate the solution orthogonally, since factors did not correlate highly with one another. Kaiser's normalization was applied to factor loadings after rotation. The principal components analysis revealed that five factors had eigenvalues that were greater than one, but inspection of a scree plot revealed an 'elbow' at the fourth factor, indicating that successive eigenvalues decreased slightly only after the fourth factor. This finding indicated that a four-factor solution was appropriate for the school-level variables. The four-factor solution accounted for approximately 58 percent of the variance in the school-level variables.

Table 4 shows that each school-level variable was well defined by the factor solution. The variables free/reduced lunch, educational climate, percent of college-bound students, relative wealth, and percent minority students loaded highly on the first factor; the variables relative wealth, metro status, number of students, percent minority students, and number of classrooms loaded on the second factor; the variables computers: total, computers: Internet, and classrooms: Internet loaded on the third factor; and the variables technology measure and innovative programs loaded highly on the fourth factor.

With the exception of the variables relative wealth and percent minority students, the loadings for each variable were high for one and only one factor. The communality values were also large for each variable with the exception of number of classrooms. The four factors were labeled economic advantage, school size, computer technology, and school resources, respectively. Since the results of the factor analysis could be clearly interpreted, the Bartlett method was used to create factor scores. These factor scores were used as level-2 independent variables. Each factor score had a mean of zero and a standard deviation of one.

---

[4] The Bartlett method uses least squares to estimate an individual's factor score over the range of variables.

Table 3
Description of School-Level Variables

School-Level Variable	Description
1. Free/Reduced Lunch	Percentage of students in institution identified as qualifying for free lunch (compensatory education) funds
2. Educational Climate	A measure of a school's socio-economic status based on its zip code; it is weighted to more strongly reflect the educational aspect of social status
3. Percent College-Bound	Percentage of graduating seniors with two and four-year college/university post-high school plans
4. Relative Wealth	Orshansky's (year) percent of universe subtracted from 100[5]
5. Computers: Total	Total number of computers in the institution
6. Computers: Internet	Total number of computer connected to the Internet
7. Classrooms: Internet	Total number of classrooms connected to the Internet
8. Metro Status	Geographic and demographic characteristic of a school ranging from 1 (large central city) to 7 (rural area).
9. Number of Students	Total number of students within an institution
10. Percent Minority	Percent of Minority (non-Caucasian) students.
11. Number of Classrooms	Total number of classrooms within an institution
12. Technology Measure[6]	School's technology presence when compared to all other schools
13. Innovative Programs	Total number of innovative programs within an institution, i.e., Advanced Placement, gifted program, etc.

Note. From Quality Education Data (2000) Data User Guide.

Table 4
Factor Loadings and Communalities for the Four-Factor Solution for School-Level Variables

	Factor Loading				
School-Level Variable	1	2	3	4	Communality
1. Free/Reduced Lunch	-0.84				0.76
2. Educational Climate	0.69				0.58
3. Percent College-Bound	0.63				0.41
4. Relative Wealth	0.65	-0.41			0.64
5. Computers: Total			0.86		0.40
6. Computers: Internet			0.84		0.76
7. Classrooms: Internet			0.59		0.72
8. Metro Status		-0.71			0.52
9. Number of Students		0.70			0.57
10. Percent Minority	0.46	0.67			0.67
11. Number of Classrooms		0.51			0.29
12. Technology Measure				0.80	0.65
13. Innovative Programs				0.70	0.52

Note: Loadings less than 0.4 are suppressed.

[5] Orshansky Percent of Universe is the number of students falling below the federal government poverty guidelines as a percentage of all children within a district's boundaries. This figure is used as a relative indicator of community wealth/poverty in comparison to other school districts (QED, 2000). QED's Relative Wealth Indicator is the Orshansky percentage subtracted from 100.

BEST COPY AVAILABLE

Results

Table 5 shows the mean student-level variables by gender and best language[7]. Mean HSGPA and curriculum intensity do not vary greatly by gender and best language group; however, SAT scores, discrepancy scores, SATM-V discrepancy, and SES factor scores are substantially different across these groups. Females have larger negative discrepancy scores than males in all three language groups; bilingual and LEP females have larger negative discrepancy scores than English-speaking females. English-speaking females score better on the SAT-V, and English-speaking males score better on the SAT-M; however, bilingual and LEP females and males score better on SAT-M. The magnitude of the difference between verbal and math scores is greater for males than for females, and is especially large for LEP students. Both males and females in the bilingual and LEP groups have substantially lower SES factor scores than English males and females.

Table 5
Means of Student-Level Variables by Gender and Language Group

Student-Level Variable	All Students		English		Bilingual		LEP	
	Females	Males	Females	Males	Females	Males	Females	Males
N	19,072	16,793	16,739	14,411	8?3	639	231	258
HSGPA	3.6	3.5	3.6	3.5	3.6	3.5	3.7	3.5
SAT-V	574	579	578	584	533	538	448	458
SAT-M	565	606	566	607	549	603	599	622
Discrepancy	-.22	.26	-.20	.28	-.47	-.02	-.89	-.46
Curriculum Intensity	14.1	14.5	14.2	14.6	14.6	15.1	14.5	14.5
SAT M-V	-8.7	27.4	-11.8	23.2	15.6	64.6	150.2	163.8
SES Factor Score	-.05	.06	-.02	.09	-.57	-.45	-.56	-.55

Figure 1 shows the interaction of gender and language group in mean discrepancy scores. LEP females and males both have negative discrepancy scores; bilingual males have a mean discrepancy score that is close to zero while bilingual females have a negative score; English

[6] Technology Measure is an indicator developed by QED that summarizes the presence of electronic technology in schools that accounts for their different sizes, student populations, and types of equipment. This variable is measured on an ordinal scale from 1 to 7 (QED, 2000).

males have a positive discrepancy score while English females have a negative score. Please note that bilingual and LEP students make up only 4.4 and 1.5 percent of the sample, respectively; the number of students in these groups is substantially smaller than in the English-speaking group.

Figure 1
Mean Discrepancy Scores by Gender and Language Groups

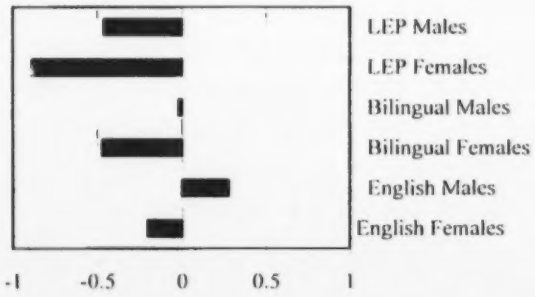


Table 6 shows the correlations of discrepancy score with curriculum intensity and SES separately for each gender and language group. The correlations for curriculum intensity and SES are positive, indicating that as curriculum becomes more intense and socio-economic conditions improve, students' SAT scores are more likely to be higher than their HSGPA. However, the relationship is stronger for females than for males, and stronger for LEP students than for English and bilingual students. Furthermore, the difference between males and females increases across each group, such that the largest difference occurs in the LEP group (See Figure 2).

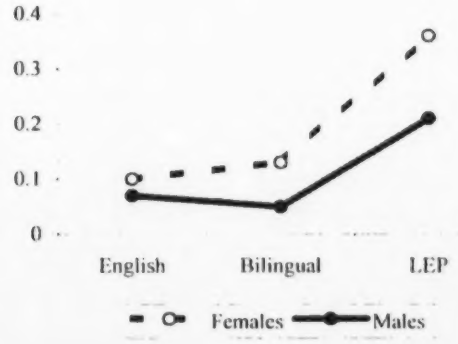
[7] Best language was measured with the following scale: 1 = English, 2 = English and another language (bilingual), 3 = Another language (LEP).

Table 6
Correlations of Curriculum Intensity and SES with Discrepancy Score
by Gender and Language Group

Variable	All Students		English		Bilingual		LEP	
	Females	Males	Females	Males	Females	Males	Females	Males
Curriculum Intensity	.10 (18,027)	.07 (15,585)	.10 (16,560)	.07 (14,247)	.13 (800)	.05 (631)	.36 (223)	.21 (250)
SES	.23 (15,475)	.21 (13,656)	.22 (14,406)	.19 (12,714)	.30 (680)	.25 (551)	.23 (192)	.38 (221)

Note. The N's are in parentheses under each correlation.

Figure 2
Correlation of Curriculum Intensity with Discrepancy Score
by Gender and Language Group



A preliminary analysis of the contribution of the student-level variables to discrepancy scores was performed using multiple regression. Table 7a shows the results of the multiple regression of discrepancy score on the student-level variables, separately for each language group. The results indicate that curriculum intensity has greater weight and gender has less weight in the model for LEP students than in the models for English or bilingual students. The SATM-V discrepancy contributed minimally to the model for LEP students, while it was a significant predictor of discrepancy scores in the English and bilingual groups.

Next, discrepancy scores were regressed on SES, curriculum intensity, gender (0 = males, 1 = females), SATM-V, LEP (1 = LEP, 0 = English or bilingual), and the two-way interaction of LEP with curriculum intensity, SES, SATM-V, and gender. Table 7b shows the standardized

coefficients and R-square values for each of the student-level variables in the regression models. The variables that were found to be significant predictors of discrepancy scores, in the order of their importance were gender, SES, LEP, SATM-V, and curriculum intensity. The interactions between LEP and curriculum intensity and SATM-V were also significant.

Table 7a
Standardized Coefficients and R-Squares from Multiple Regression of Discrepancy Score on Student-Level Variables

Student-Level Variable	All Students		English		Bilingual		LEP	
	B	R^2	B	R^2	B	R^2	B	R^2
SES	.20*	.051	.19*	.045	.25*	.078	.24*	.092
INTENSE	.07*	.056	.07*	.050	.08*	.083	.23*	.141
Gender	-.25*	.106	-.25*	.101	-.23*	.123	-.18*	.172
SAT M-V	-.11*	.117	-.10*	.110	-.13*	.137	-.01	.172

Note. * $p < .05$

Table 7b
Standardized Coefficients (Beta) and R-Squares from Multiple Regression of Discrepancy Score on Student-Level Variables

Student-Level Variable	Beta	R-Square	Significance
SES	.20	.051	.000
INTENSE	.07	.056	.000
Gender	-.25	.106	.000
LEP	-.19	.110	.000
SAT M-V	-.11	.119	.000
LEP*INTENSE	.11	.120	.000
LEP*SES	.01	.120	.423
LEP*SAT M-V	.03	.120	.000
LEP*Gender	.02	.120	.235

Hierarchical Linear Model

The exploratory analyses described above guided the choice of student-level variables to include in the hierarchical linear modeling of the influence of student and school characteristics on discrepancy scores. The variance components model was the first model fit to the data. The level-1 equation for this model is:

(1) $D_{ij} = \beta_{0j} + r_{ij}$

This equation represents each student's discrepancy score (D_{ij}) as the mean within his or her school plus a residual. The level-2 equation for this model is:

(2) $\beta_{0j} = \gamma_{00} + \mu_{0j}$,

which represents each school's average discrepancy score as an average for all schools, plus a residual for each school.

The results of the variance components model showed that the average discrepancy score was approximately zero (-0.01). The variance across schools in the average discrepancy score was 0.22 (SD = .47). Thus school means would be expected to vary from approximately –0.95 to 0.93. The variance of discrepancy scores among students within a school is 0.78 (SD = .89). Therefore, student scores would be expected to vary from the school mean up or down about 1.77 points. One can see from these results that there is significant variation in discrepancy scores both within and between schools. Student and school-level predictors of these sources of variability were explored in subsequent models.

The second model fit to the data extended the analysis to include the student-level variables deemed important during preliminary analysis: curriculum intensity (INTENSE), socio-economic status (SES), gender, and SAT M-V. Student-level variables were not centered during estimation of this and subsequent models. The LEP variable was not included in the model even though it was found to be an important predictor of discrepancy scores, because a substantial number of schools did not have enough variability on this variable to be included in the analysis. (When included, the number of schools that could be analyzed by HLM 5 dropped from 781 to 126). The level-1 equation for this model is:

(3) $D_{ij} = \beta_{0j} + \beta_{1j}(curr_{ij}) + \beta_{2j}(ses_{ij}) + \beta_{3j}(gender_{ij}) + \beta_{4j}(M - V_{ij}) + r_{ij}$,

which models a student's discrepancy score as a function of his or her school mean and score on the student level variables considered in the analysis. There are five level-2 equations associated with this model, one for each level one coefficient:

(4) $\beta_{0j} = \gamma_{00} + \mu_{0j}$

(5) $\beta_{1j} = \gamma_{10} + \mu_{1j}$

(6) $\beta_{2j} = \gamma_{20} + \mu_{2j}$

(7) $\beta_{3j} = \gamma_{30} + \mu_{3j}$

(8) $\beta_{4j} = \gamma_{40} + \mu_{4j}$

These equations allow both the school intercepts and slopes to vary across schools without any school-level predictors.

The results of this model show that all of the student-level variables except for curriculum intensity are statistically significant predictors of discrepancy scores (see Table 8). Level-one intercepts as well as slopes for gender varied significantly across schools; the slopes for the other student-level variables did not. This finding warranted exploring school-level predictors of those level-one coefficients that varied significantly across schools. The results also indicated that the variance of the residual for the level one equation changed from 0.78 in the first model (which had no student-level predictors) to 0.70 in the second model (which had four student-level predictors). This result suggests that while several of student-level variables are statistically significant predictors of discrepancy scores, much of the variation in these scores remains unexplained.

Additional models were fit to the data to predict the variability in the level-one intercepts and slopes for gender across schools. The results indicated that the factor scores for economic advantage and school size significantly predicted a school's average discrepancy score (or level

one intercept) ($t_{(777)} = 12.37$, $p < .05$ and $t_{(777)} = 5.08$, $p < .05$, respectively) The coefficients and standard errors for the economic advantage and school size factors were 0.19 (0.01) and 0.08 (0.01), respectively. The economic advantage factor also predicted a school's slope for gender ($t_{(777)} = 5.9$, $p < .05$); the coefficient was 0.08 (0.01). The variance component of the random effect for level-one intercepts decreased by only .01 (from 0.36 to 0.35), even after economic advantage and school size were considered in the analysis, indicating that significant variability remains unexplained for the level-one intercepts. The variance component for the slopes for gender (.03) remained the same after economic advantage was considered in the analysis.

Table 8
Hierarchical Linear Model Results with Level-1 Predictors

Level 1					
Fixed Effect	Standardized Coefficient	Standard Error	Approximate T-ratio	df	P-value
Intercept	0.24	0.03	7.44	780	0.00
Curr	0.00	0.00	1.17	780	0.24
SES	0.09	0.01	12.47	780	0.00
Gender	-0.49	0.02	-31.54	780	0.00
SAT M-V	-0.00	0.00	-17.29	780	0.00
Level 2					
Random Effect	Standard Deviation	Variance Component	Chi-square	df	P-value
Intercept	0.60	0.36	1,203.71	777	0.00
Slope: curr	0.01	0.00	770.99	777	0.16
Slope: SES	0.02	0.00	690.37	777	>0.50
Slope: Gender	0.16	0.03	801.14	777	0.04
Slope: SAT M-V	0.00	0.00	780.87	777	0.11
Level 1 residual	0.84	0.70	---	---	---

Note. The chi-square statistics reported above are based on only 778 of 781 units that had sufficient data for computation. Fixed effects and variance components are based on all the data.

Discussion

This study attempted to explain discrepancies between students' high school grades and SAT scores as a function of both student-level and school-level variables. As found by Kobrin and Milewski (2002), an individual's gender and best language are strong predictors of the discrepancy between HSGPA and SAT scores. Females, and students for whom English is not their best language tend to have higher high school grades than SAT scores, while English-

speaking males tend to have higher SAT scores than high school grades. Preliminary regression analyses indicated that there are different relationships between curriculum intensity, SES, and SATM-V for different best language groups. However, not enough of the schools in this study had a large enough population of LEP students to further explore these relationships using HLM.

Other student-level variables that were found to predict the discrepancy between high school grades and SAT scores were socio-economic status and the difference between SAT Math and Verbal scores. Students with higher SES were more likely to have SAT scores that were in line with, or higher than, their high school grades. It is important to remember that the factor score for SES took into account both parental education and family income, which may reflect the influence of a student's educational climate in the home. This suggests that students from more affluent and better-educated family backgrounds are advantaged by the virtue of having backgrounds and educational experiences that foster the development of academic abilities assessed by the SAT.

Students who obtained higher SAT verbal scores relative to their math scores also tended to have higher high school grades than their total SAT scores. One might surmise that students who obtain high scores on the verbal test also tend to take English and humanities courses rather than math and science courses in high school, and that the former courses are typically graded easier than the latter, enabling these students to earn higher grades relative to their SAT scores. However, the findings of this study suggest that the intensity of a student's curriculum in high school does not predict the likelihood of having discrepant SAT scores and high school grades. There is one caveat to note, however, course-taking patterns were self-reported and may be unreliable. In addition, many of the students in this study had missing data on courses taken, making it difficult to measure curriculum intensity for these students. Finally, the data were

based on students who had completed one year of college and had a valid grade point average. These issues may have biased the results of our analyses.

While several of the student-level variables were significant predictors of discrepancy scores, a substantial amount of the variance remains unexplained. This suggests that other variables not examined in this study are important predictors of the discrepancy between high school grades and SAT scores. These variables may include student motivation and teacher grading standards, two variables that are very difficult to measure in large-scale studies.

This study found that most of the variability of discrepancy scores was within schools rather than between schools. However, the intercepts (a school's average discrepancy score) in the level-one models did vary significantly across schools, suggesting that there are certain school characteristics that influence the discrepancies. The relationship of gender and discrepancy scores also varied significantly across schools, while the slopes for the remaining school-level variables were essentially uniform across schools. Four school-level factors were used to predict the variability in the school intercepts and slopes for gender: economic advantage, school size, computer technology, and school resources. The economic advantage of a school, which encompassed the school's percentage of students eligible for free/reduced lunch, educational climate, percent of college-bound students, and relative wealth; and school size were found to be significant predictors of the schools' average discrepancy score. The economic advantage of a school also significantly predicted the relationship between gender and discrepancy scores across schools.

These findings suggest that a school's economic condition is an important factor in the prevalence of discrepant HSGPA and SAT scores. If a school has good economic conditions, students may be less likely to have a discrepancy between their high school grades and SAT scores, and there may not be as big a difference between females and males in the frequency of

this discrepancy. This finding confirms a long-standing belief that students from schools in low-income areas are not as academically prepared as their higher income counterparts. It is further evidence of unequal access to quality education. These findings also raise questions about grade inflation, an area of educational measurement that is not well understood. If students from economically disadvantaged schools with poor quality educational programs that do not prepare them for college are earning the same grades as students from higher quality schools, then we ask are their grades artificially inflated? This is a question that deserves more attention.

Previous research has suggested that students with a high HSGPA in the presence of low SAT scores will not do any better in college than students with lower HSGPA scores but higher SAT scores. Therefore, the SAT may be a more accurate predictor than HSGPA for the former students. This study identified some student and school characteristics that are important predictors of discrepant scores, however, much work remains to be done to elucidate this phenomenon so that students, teachers, guidance counselors and college admissions staff can better understand the relationship between these two indicators of student performance. Future research will focus on identifying and measuring other student- and school-level variables that might influence the discrepancy, such as grade inflation, teacher grading standards, and student motivation. Future research will also be based on a sample of students that is more representative of the college-bound population.

References

Adelman, C. (1999). *Answers in the tool box: Academic intensity, attendance patterns, and bachelor's degree attainment*. Washington, DC: U.S. Department of Education.

Baydar, N. (1990). *Profiles of the students who have discrepant high school GPA and SAT scores*. Unpublished manuscript. Princeton, NJ: Educational Testing Service.

Bejar, I.I., & Blew, E.O. (1981). *Grade inflation and the validity of the Scholastic Aptitude Test* (College Board Rep. No. 81-3). New York, NY: College Entrance Examination Board.

Breland, H., Maxey, J., Gernand, R., Cumming, T., & Trapani, C. (2002). *Trends in college admission 2000: A report of a survey of undergraduate admission policies, practices and procedures*. (ACT, AIR, College Board, ETS, NACAC).

Bridgeman, B., McCamley-Jenkins, L., & Ervin, N. (2000). *Predictions of freshman grade-point average from the revised and recentered SAT I: Reasoning test*. (College Board Report No. 2000-1). New York: College Entrance Examination Board.

Bryk, A. S., & Raudenbush, S. W. (1992). *Hierarchical linear models: Applications and data analysis methods*. Newbury Park: Sage.

Cizek, G.J. (2000). Pockets of resistance in the assessment revolution. *Educational Measurement: Issues and Practice, 19*, 16-23,33.

College Board. (2001). *Register for the SAT*. New York: The College Board.

Everson, H. & Millsap, R. (1999). *Predictors of performance on the SAT: Multilevel models of individual differences and school effects*. Paper presented at the NCAA DARN Meeting (San Diego, CA, December 10-11, 1999).

Kobrin, J.L., & Milewski, G.B. (2002). *Students with discrepant high school GPA and SAT I scores*. (College Board Research Note No. RN-15). New York: College Entrance Examination Board.

Koretz, D., & Berends, M. (2001). *Changes in high school grading standards in mathematics, 1982-1992*. (RAND Research Report MR-1445-CB). Available at http://www.rand.org/publications/MR/MR1445.

McCullagh, P., & Nelder, J. (1989). *Generalized linear models* (2nd edition). London: Chapman and Hill.

Mulvenon, S.V., & Thorn, A.R. (2002). *Remediation rates and grade point averages: Is there grade inflation in Arkansas?* Unpublished manuscript.

National Center of Education Statistics (NCES). http://nces.ed.gov/ccd/pubschuniv.html.

Pilcher, J.K. (1994). The value-driven meaning of grades. *Educational Assessment, 2*(1), 69-88.

Quality Education Data (2000). *National Education Database Data User Guide, Version 4.8.* Denver, CO: Quality Education Data.

Raudenbush, S., Bryk, A., Cheong, Y. F., Congdon, R. (2000). *HLM 5: Hierarchical linear and nonlinear modeling.* Chicago, IL: Scientific Software International.

Stiggins, R.J., Frisbie, D.A., & Griswold, P.A. (1989). Inside high school grading practices: Building a research agenda. *Educational Measurement: Issues and Practice,* Summer 1989, 5-14.

U.S. Department of Education, Office of Educational Research and Improvement (1994). *What do student grades mean? Differences across schools.* Washington, DC: U.S. Department of Education. Available at http://www.ed.gov/pubs/OR/ResearchRpts/grades.html.

Vogt, P. V. (1999). *Dictionary of statistics & methodology: A nontechnical guide for the social sciences* (2nd edition). Thousand Oaks: Sage.

Willingham, W.W., Lewis, C., Morgan, R., & Ramist, L. (1990). *Predicting college grades: An analysis of institutional trends over two decades.* Princeton, NJ: Educational Testing Service.

Ziomek, R.L., & Svec, J.C. (1995). *High school grades and achievement: Evidence of grade inflation.* (ACT Research Report Series 95-3). Iowa City, IA: The American College Testing Program.

Zirkel, P.A. (1999). Grade inflation: A leadership opportunity for schools of education? *Teachers College Record, 101,* 247-260.

U.S. Department of Education
Office of Educational Research and Improvement (OERI)
National Library of Education (NLE)
Educational Resources Information Center (ERIC)

REPRODUCTION RELEASE

(Specific Document)

TM034972

I. DOCUMENT IDENTIFICATION:

Title: An Investigation of School-Level Factors for Students with Discrepant High School GPA and SAT Scores	
Author(s): Jennifer L. Kobrin, Glenn B. Milewski, Howard Everson, and Ying Zhou	
Corporate Source: The College Board	Publication Date: April 22, 2003

II. REPRODUCTION RELEASE:

In order to disseminate as widely as possible timely and significant materials of interest to the educational community, documents announced in the monthly abstract journal of the ERIC system, *Resources in Education* (RIE), are usually made available to users in microfiche, reproduced paper copy, and electronic media, and sold through the ERIC Document Reproduction Service (EDRS). Credit is given to the source of each document, and, if reproduction release is granted, one of the following notices is affixed to the document.

If permission is granted to reproduce and disseminate the identified document, please CHECK ONE of the following three options and sign at the bottom of the page.

The sample sticker shown below will be affixed to all Level 1 documents

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY

Sample

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

1

Level 1

☑

Check here for Level 1 release, permitting reproduction and dissemination in microfiche or other ERIC archival media (e.g., electronic) and paper copy.

The sample sticker shown below will be affixed to all Level 2A documents

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL IN MICROFICHE, AND IN ELECTRONIC MEDIA FOR ERIC COLLECTION SUBSCRIBERS ONLY, HAS BEEN GRANTED BY

Sample

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

2A

Level 2A

☐

Check here for Level 2A release, permitting reproduction and dissemination in microfiche and in electronic media for ERIC archival collection subscribers only

The sample sticker shown below will be affixed to all Level 2B documents

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL IN MICROFICHE ONLY HAS BEEN GRANTED BY

Sample

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

2B

Level 2B

☐

Check here for Level 2B release, permitting reproduction and dissemination in microfiche only

Documents will be processed as indicated provided reproduction quality permits.
If permission to reproduce is granted, but no box is checked, documents will be processed at Level 1.

I hereby grant to the Educational Resources Information Center (ERIC) nonexclusive permission to reproduce and disseminate this document as indicated above. Reproduction from the ERIC microfiche or electronic media by persons other than ERIC employees and its system contractors requires permission from the copyright holder. Exception is made for non-profit reproduction by libraries and other service agencies to satisfy information needs of educators in response to discrete inquiries.

Sign here, → please

Signature: Jennifer L. Kobrin	Printed Name/Position/Title: Jennifer L. Kobrin, Associate Research Scien
Organization/Address: The College Board, 45 Columbus Avenue, New York, NY 10023-6992	Telephone: 212-713-8318 FAX: 212-649-8427
	E-Mail Address: jkobrin@collegeboard.org Date: 5/21/03

(Over)

III. DOCUMENT AVAILABILITY INFORMATION (FROM NON-ERIC SOURCE):

If permission to reproduce is not granted to ERIC, *or*, if you wish ERIC to cite the availability of the document from another source, please provide the following information regarding the availability of the document. (ERIC will not announce a document unless it is publicly available, and a dependable source can be specified. Contributors should also be aware that ERIC selection criteria are significantly more stringent for documents that cannot be made available through EDRS.)

Publisher/Distributor:
Address:
Price:

IV. REFERRAL OF ERIC TO COPYRIGHT/REPRODUCTION RIGHTS HOLDER:

If the right to grant this reproduction release is held by someone other than the addressee, please provide the appropriate name and address:

Name:
Address:

V. WHERE TO SEND THIS FORM:

Send this form to the following ERIC Clearinghouse:

University of Maryland
ERIC Clearinghouse on Assessment and Evaluation
1129 Shriver Lab, Bldg 075
College Park, MD 20742
Attn: Acquisitions

However, if solicited by the ERIC Facility, or if making an unsolicited contribution to ERIC, return this form (and the document being contributed) to:

University of Maryland
ERIC Clearinghouse on Assessment and Evaluation
1129 Shriver Lab, Bldg 075
College Park, MD 20742
Attn: Acquisitions